

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
"Казанский (Приволжский) федеральный университет"
Институт вычислительной математики и информационных технологий



УТВЕРЖДАЮ

Проректор по образовательной деятельности КФУ
проф. Таюрский Д.А.

"__" _____ 20__ г.

Программа дисциплины

Программирование в среде R

Направление подготовки: 01.04.02 - Прикладная математика и информатика

Профиль подготовки: Методы прикладной математической статистики

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО
2. Место дисциплины (модуля) в структуре ОПОП ВО
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся
4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий
 - 4.1. Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)
 - 4.2. Содержание дисциплины (модуля)
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)
6. Фонд оценочных средств по дисциплине (модулю)
7. Перечень литературы, необходимой для освоения дисциплины (модуля)
8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)
12. Средства адаптации преподавания дисциплины (модуля) к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья
13. Приложение №1. Фонд оценочных средств
14. Приложение №2. Перечень литературы, необходимой для освоения дисциплины (модуля)
15. Приложение №3. Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Программу дисциплины разработал(а)(и) доцент, к.н. (доцент) Григорьева И.С. (кафедра математической статистики, отделение прикладной математики и информатики), Irina.Grigorieva@kpfu.ru

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО

Обучающийся, освоивший дисциплину (модуль), должен обладать следующими компетенциями:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОК-1	способность к абстрактному мышлению, анализу, синтезу
ОК-3	готовность к саморазвитию, самореализации, использованию творческого потенциала
ОПК-3	способность самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе, в новых областях знаний, непосредственно не связанных со сферой деятельности, расширять и углублять своё научное мировоззрение
ОПК-4	способность использовать и применять углубленные знания в области прикладной математики и информатики
ПК-8	способность разрабатывать корпоративные стандарты и профили функциональной стандартизации приложений, систем, информационной инфраструктуры
ПК-9	способность к преподаванию математических дисциплин и информатики в образовательных организациях основного общего, среднего общего, среднего профессионального и высшего образования

Обучающийся, освоивший дисциплину (модуль):

Должен демонстрировать способность и готовность:

Применять язык и среду R для написания программ, позволяющих решать основные задачи статистической обработки данных

Самостоятельно изучать новые возможности языка

Использовать сведения разработчиков R, а также пользователей для расширения знаний о среде R

Самостоятельно составлять комплексы программ для решения прикладных задач статистики и анализа данных

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.ОД.2 Дисциплины (модули)" основной профессиональной образовательной программы 01.04.02 "Прикладная математика и информатика (Методы прикладной математической статистики)" и относится к обязательным дисциплинам.

Осваивается на 1 курсе в 1 семестре.

3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Общая трудоемкость дисциплины составляет 2 зачетных(ые) единиц(ы) на 72 часа(ов).

Контактная работа - 28 часа(ов), в том числе лекции - 0 часа(ов), практические занятия - 0 часа(ов), лабораторные работы - 28 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 44 часа(ов).

Контроль (зачёт / экзамен) - 0 часа(ов).

Форма промежуточного контроля дисциплины: зачет в 1 семестре.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1 Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Основные объекты языка R (векторы, факторы, таблицы и списки). Векторные вычисления. Простейшие графические команды. Ввод данных из внешнего источника. Создание простых скриптов	1	0	0	6	10
2.	Тема 2. Датчики случайных величин с разным распределением. Основные статистические задачи (вычисление параметров, доверительных интервалов, построение гистограмм, проверка гипотез) Создание банка программ для обработки табличных данных	1	0	0	6	8
3.	Тема 3. Подключение библиотек. Методы проверки нормальности данных. Библиотека анализа мощности. Библиотека трехмерной визуализации	1	0	0	4	8
4.	Тема 4. Группа методов построения линейных моделей. Команды <code>lm()</code> , <code>glm()</code> . Построение линий регрессии. Логистическая регрессия	1	0	0	4	8
5.	Тема 5. Кластеризация данных. Методы k-means и иерархической кластеризации. Визуализация результатов кластеризации. 3d-объекты.	1	0	0	6	8
6.	Тема 6. Применение полученных навыков для самостоятельного анализа данных	1	0	0	2	2
	Итого		0	0	28	44

4.2 Содержание дисциплины (модуля)

Тема 1. Основные объекты языка R (векторы, факторы, таблицы и списки). Векторные вычисления. Простейшие графические команды. Ввод данных из внешнего источника. Создание простых скриптов

Повторение основных сведений о целях, особенностях, структуре и объектах статистической среды R. Вектор - базовый объект языка. Векторные вычисления - специфический метод организации работы с данными. Многомерные объекты (массивы, матрицы, таблицы), их особенности. Способы извлечения данных из многомерных объектов. Фактор - объект для обработки категориальных данных. Использование прямых команд для вычислений. Использование команд группы `apply()`. Простейшие графические операторы (`plot`, `lines`, `points`, `abline` и т.п.) Построение графиков и скаттерплов. Первичная визуализация данных. Ввод данных из текстовых файлов и файлов типа `.csv`. Проверка результатов ввода, автоматический поиск ошибок. Создание скриптов, запуск их из консоли. Особенности синтаксиса команд на экране и в файле.

Тема 2. Датчики случайных величин с разным распределением. Основные статистические задачи (вычисление параметров, доверительных интервалов, построение гистограмм, проверка гипотез) Создание банка программ для обработки табличных данных

Создание модельных объектов с помощью датчиков случайных чисел, а также команд группы `as.*()` Вычисление основных параметров: среднее, дисперсия, стандартное отклонение, медиана, квантили, коэффициенты корреляции. Особенности применения соответствующих команд к данным разных типов. Доверительные интервалы для параметров. Проверка гипотез: команды группы `*.test()`. Структура входной и выходной информации. Создание собственных списков вывода. Создание подпрограммы для запуска произвольной команды проверки гипотезы и обработки результата. Вывод результатов в текстовый файл. Возможности языка R: использование имени команды в качестве фактического параметра процедуры. Использование текстовых строк для создания имен выводимых файлов.

Тема 3. Подключение библиотек. Методы проверки нормальности данных. Библиотека анализа мощности. Библиотека трехмерной визуализации

Библиотеки в R. Знакомство с набором библиотек по литературе и сайтам, в том числе официальному сайту R. Возможности по преобразованию данных. Библиотека проверки нормальности. Создание тестовых наборов данных для проверки. Создание скриптов, применяющих различные методы проверки к данным. Сравнение результатов такой проверки.

Библиотека анализа мощности `powr`. Анализ мощности теста Стьюдента, кор-теста. Подбор размера выборки по желаемым вероятностям ошибок первого и второго рода.

Библиотека `rgl`. Построение интерактивных 3d-графиков.

Тема 4. Группа методов построения линейных моделей. Команды `lm()`, `glm()`. Построение линий регрессии. Логистическая регрессия

Команда `lm()` и создаваемый ею объект. Использование полученного объекта для поиска зависимостей в данных. Команда `glm()`. Использование категориальных переменных в качестве отклика и независимых переменных. Построение линий регрессии (линейной и нелинейной).

Визуализация многомерных данных. Методы главных компонент.

Тема 5. Кластеризация данных. Методы k-means и иерархической кластеризации. Визуализация результатов кластеризации. 3d-объекты.

Изучение различных способов кластеризации данных. Создание тестовых (двумерных) наборов данных в различной структуре. Метод k-means разбиения на фиксированное число классов. Тестирование метода на заготовленных тестовых наборах. Влияние случайности и эвристичности метода. Метод иерархической кластеризации. Подготовка данных (вычисление матрицы расстояний командой `dist()`). Применение метода и визуализация полученного дерева. Выделение классов. Визуализация классов на скаттерплоте с использованием цвета и значков разного типа. Визуализация многомерных данных с помощью метода главных компонент.

Сравнение различных вариантов методов кластеризации.

Тема 6. Применение полученных навыков для самостоятельного анализа данных

Самостоятельное создание скриптов, их отладка и прогонка на тестовых данных.

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Самостоятельная работа обучающихся выполняется по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Самостоятельная работа подразделяется на самостоятельную работу на аудиторных занятиях и на внеаудиторную самостоятельную работу. Самостоятельная работа обучающихся включает как полностью самостоятельное освоение отдельных тем (разделов) дисциплины, так и проработку тем (разделов), осваиваемых во время аудиторной работы. Во время самостоятельной работы обучающиеся читают и конспектируют учебную, научную и справочную литературу, выполняют задания, направленные на закрепление знаний и отработку умений и навыков, готовятся к текущему и промежуточному контролю по дисциплине.

Организация самостоятельной работы обучающихся регламентируется нормативными документами, учебно-методической литературой и электронными образовательными ресурсами, включая:

Порядок организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры (утвержден приказом Министерства образования и науки Российской Федерации от 5 апреля 2017 года №301)

Письмо Министерства образования Российской Федерации №14-55-996ин/15 от 27 ноября 2002 г. "Об активизации самостоятельной работы студентов высших учебных заведений"

Устав федерального государственного автономного образовательного учреждения "Казанский (Приволжский) федеральный университет"

Правила внутреннего распорядка федерального государственного автономного образовательного учреждения высшего профессионального образования "Казанский (Приволжский) федеральный университет"

Локальные нормативные акты Казанского (Приволжского) федерального университета

Основной сайт языка R - <http://cran.r-project.org/>

Советы по использованию R - <http://pj.freefaculty.org/R/Rtips.html>

6. Фонд оценочных средств по дисциплине (модулю)

Фонд оценочных средств по дисциплине (модулю) включает оценочные материалы, направленные на проверку освоения компетенций, в том числе знаний, умений и навыков. Фонд оценочных средств включает оценочные средства текущего контроля и оценочные средства промежуточной аттестации.

В фонде оценочных средств содержится следующая информация:

- соответствие компетенций планируемым результатам обучения по дисциплине (модулю);
- критерии оценивания сформированности компетенций;

- механизм формирования оценки по дисциплине (модулю);
- описание порядка применения и процедуры оценивания для каждого оценочного средства;
- критерии оценивания для каждого оценочного средства;
- содержание оценочных средств, включая требования, предъявляемые к действиям обучающихся, демонстрируемым результатам, задания различных типов.

Фонд оценочных средств по дисциплине находится в Приложении 1 к программе дисциплины (модулю).

7. Перечень литературы, необходимой для освоения дисциплины (модуля)

Освоение дисциплины (модуля) предполагает изучение основной и дополнительной учебной литературы. Литература может быть доступна обучающимся в одном из двух вариантов (либо в обоих из них):

- в электронном виде - через электронные библиотечные системы на основании заключенных КФУ договоров с правообладателями;

- в печатном виде - в Научной библиотеке им. Н.И. Лобачевского. Обучающиеся получают учебную литературу на абонементе по читательским билетам в соответствии с правилами пользования Научной библиотекой.

Электронные издания доступны дистанционно из любой точки при введении обучающимся своего логина и пароля от личного кабинета в системе "Электронный университет". При использовании печатных изданий библиотечный фонд должен быть укомплектован ими из расчета не менее 0,5 экземпляра (для обучающихся по ФГОС 3++ - не менее 0,25 экземпляра) каждого из изданий основной литературы и не менее 0,25 экземпляра дополнительной литературы на каждого обучающегося из числа лиц, одновременно осваивающих данную дисциплину.

Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля), находится в Приложении 2 к рабочей программе дисциплины. Он подлежит обновлению при изменении условий договоров КФУ с правообладателями электронных изданий и при изменении комплектования фондов Научной библиотеки КФУ.

8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

CRAN - <http://cran.r-project.org/>

R-help - <https://stat.ethz.ch/pipermail/r-help>

Советы по использованию - http://zoonek2.free.fr/UNIX/48_R/all.html

Советы по использованию - <http://www.statmethods.net/index.html>

Советы по использованию R - <http://pj.freefaculty.org/R/Rtips.html>

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Язык R -- современный быстро развивающийся язык программирования, применяющийся для решения статистических задач.

Для успешного освоения необходимо:

- знать (вспомнить, изучить) основные методы статистики (оценка параметров, проверка гипотез, дисперсионный анализ, построение линий регрессии и других моделей)
- составить и постоянно пополнять свой список интернет-ресурсов, посвященных языку R.
- активно пользоваться справочной системой языка.
- установить на домашнем компьютере среду программирования R и иметь копии всех составленных на занятиях программ.
- оптимизировать и обобщать программы, составленные на занятиях.

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем, представлен в Приложении 3 к рабочей программе дисциплины (модуля).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Материально-техническое обеспечение образовательного процесса по дисциплине (модулю) включает в себя следующие компоненты:

Помещения для самостоятельной работы обучающихся, укомплектованные специализированной мебелью (столы и стулья) и оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду КФУ.

Учебные аудитории для контактной работы с преподавателем, укомплектованные специализированной мебелью (столы и стулья).

Компьютер и принтер для распечатки раздаточных материалов.

Компьютерный класс.

12. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья

При необходимости в образовательном процессе применяются следующие методы и технологии, облегчающие восприятие информации обучающимися инвалидами и лицами с ограниченными возможностями здоровья:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников - например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения - аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной, за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;
- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи:
- продолжительности сдачи зачёта или экзамена, проводимого в письменной форме, - не более чем на 90 минут;
- продолжительности подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, - не более чем на 20 минут;
- продолжительности выступления обучающегося при защите курсовой работы - не более чем на 15 минут.

Программа составлена в соответствии с требованиями ФГОС ВО и учебным планом по направлению 01.04.02 "Прикладная математика и информатика" и магистерской программе "Методы прикладной математической статистики".

Приложение 2
к рабочей программе дисциплины (модуля)
Б1.В.ОД.2 Программирование в среде R

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 01.04.02 - Прикладная математика и информатика

Профиль подготовки: Методы прикладной математической статистики

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

Основная литература:

1. Лагутин, М.Б. Наглядная математическая статистика [Электронный ресурс] : учеб. пособие ? Электрон. дан. ? Москва : Издательство 'Лаборатория знаний', 2015. ? 475 с. ? Режим доступа: <https://e.lanbook.com/book/70706>

2. Свешников А.А. Прикладные методы теории вероятностей. - М.: Лань, 2012. - 480 с.
ЭБС 'Лань': http://e.lanbook.com/books/element.php?pl1_id=3184

3. Алон Н., Спенсер Дж. Вероятностный метод. - М.: БИНОМ. Лаборатория знаний, 2015. - 323 с.
ЭБС 'Лань': <http://e.lanbook.com/view/book/70705/>

Дополнительная литература:

1. А.Б.Шипунов, Е.М.Балдин, П.А.Волкова, А.И.Коробейников, С.А.Назарова, С.В.Петров, В.Г.Суфиянов. Наглядная статистика. Используем R! / Москва, ДМК-Пресс, 2012.- 298 с.

2. Теория вероятностей и математическая статистика: Учебник / Е.С. Кочетков, С.О. Смерчинская, В.В. Соколов. - 2-е изд., испр. и перераб. - М.: Форум: НИЦ ИНФРА-М, 2014. - 240 с.
ЭБС 'ZNANIUM.COM': <http://znanium.com/bookread.php?book=447828>

Приложение 3
к рабочей программе дисциплины (модуля)
Б1.В.ОД.2 Программирование в среде R

Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Направление подготовки: 01.04.02 - Прикладная математика и информатика

Профиль подготовки: Методы прикладной математической статистики

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

Освоение дисциплины (модуля) предполагает использование следующего программного обеспечения и информационно-справочных систем:

Операционная система Microsoft Windows 7 Профессиональная или Windows XP (Volume License)

Пакет офисного программного обеспечения Microsoft Office 365 или Microsoft Office Professional plus 2010

Браузер Mozilla Firefox

Браузер Google Chrome

Adobe Reader XI или Adobe Acrobat Reader DC

Kaspersky Endpoint Security для Windows

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен обучающимся. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, учебно-методические комплексы, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего образования (ФГОС ВО) нового поколения.